



0 651 329 A2

EUROPEAN PATENT APPLICATION

(21) Application number: 94116900.5

⑤ Int. Cl.⁶: G06F 9/46

②② Date of filing: 26.10.94

③ Priority: 28.10.93 US 144280

④3 Date of publication of application:
03.05.95 Bulletin 95/18

⑧ Designated Contracting States:
DE FR GB

71 Applicant: **MICROSOFT CORPORATION**
One Microsoft Way
Redmond,
Washington 98052-6399 (US)

(72) Inventor: **Ogus, Aaron William**
9421 112th Avenue Northeast
Kirkland,
Washington 99033 (US)

**74 Representative: Patentanwälte Grünecker,
Kinkeldey, Stockmair & Partner
Maximilianstrasse 58
D-80538 München (DE)**

⑤4 Frame buffering of network packets.

57) An improved buffering mechanism for network packets is provided. In a preferred embodiment of the present invention, a circular buffer is used when sending and receiving network packets. The circular buffer is accessible by V86 mode network drivers and protected mode applications. Said circular buffer stores packets in a first-in-first-out manner, facilitates asynchronous communications across the network and services multiple clients.

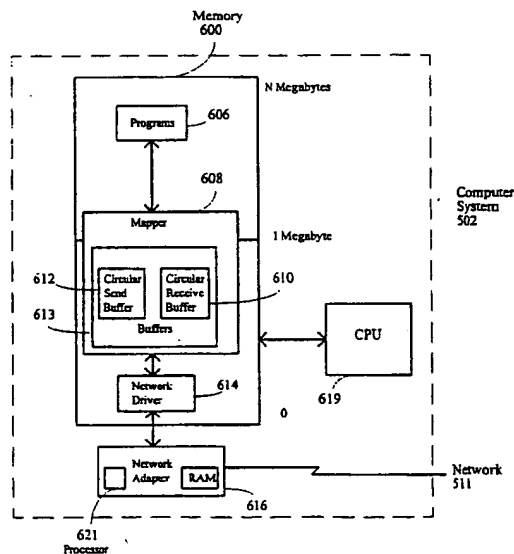


Figure 6

Field of the Invention

The present invention relates generally to data processing systems and, more particularly, to the buffering of packets sent to or received from a network.

Background of the Invention

Many modem microprocessors are part of an evolving product line of microprocessors provided by their respective manufacturers. The 80386 microprocessor, sold by Intel Corporation of Cupertino, California, is an example of such a microprocessor. The 80386 microprocessor evolved from the 80286 microprocessor, which, in turn, evolved from the 8086 and the 8088 microprocessors. The later developed microprocessors of the 80x86 family have been designed to be compatible with their earlier developed counterparts and the 8088. Hence, the 80386 microprocessor can run programs that are written for the 8086 or the 8088 microprocessor (i.e., the Intel microprocessors exhibit upward compatibility).

The 8088 microprocessor is designed to run in a single mode of operation, known as "real mode." Figure 1 depicts the addressing scheme that is employed by the 8088 microprocessor in real mode. In real mode, a 16 bit segment 100 and a 16 bit offset 102 are combined to form a 20 bit physical address 104 that specifies a location within a physical address space. The physical address 104 is formed by shifting the segment 100 four bits to the left and then adding the segment 100 with the offset 102 as shown in Figure 1. The result is a physical address 104 that identifies a memory location 110 within a segment 108 of physical memory 106. Since 20 bits are used for physical addresses in the 8088 microprocessor, the microprocessor can only address one megabyte of memory (i.e., 2^{20} Bytes). This limitation is due to the use of a 20 bit address bus in the microprocessor.

The 80386 microprocessor has the ability to run in three modes: real mode, protected mode and virtual 8086 mode (also known as V86 mode). The 80386 real mode of operation is sufficiently analogous to that of the 8088 microprocessor so as to facilitate upward compatibility with programs written on the 8088 microprocessor. That is, computer programs written for the 8088 microprocessor can execute in V86 mode on the 80386 microprocessor. Therefore, programs executing in V86 mode on the 80386 microprocessor can only access the first megabyte of memory.

The 80386 protected mode of operation provides for both segmentation and paging. Segmentation divides the physical memory into variable length blocks of memory called segments. Ad-

ressing is then done in terms of an offset within a segment instead of a direct physical address. Paging is a method of mapping a page from a linear address space into a page frame in the physical address space. The pages of memory are 4k in size. The combined use of segmentation and paging in protected mode on the 80386 microprocessor supports virtualization of the address space. Virtualization refers to a technique that allows a program to view the computer system as having a larger physical memory than the computer system actually has. In a virtualized computer system, programs access physical memory by specifying an address known as a virtual address that is then translated into a physical address through the use of paging and segmentation. Therefore, when the 80386 microprocessor is executing in protected mode, the address specified by a computer program is translated into a physical address location in a different manner than when a program is executing in real mode.

Figure 2 provides a block diagram of the addressing scheme employed in 80386 protected mode. In 80386 protected mode, the 80386 microprocessor employs a 48-bit virtual address 200. Bits 0-31 of the virtual address 200 hold an offset value 202 and bits 32-47 hold a selector 204 that is used as an index for an entry 206 in a table of addresses of segments, called a segment descriptor table 208. The entry 206 in the segment descriptor table 208 holds the base address of a segment. The base address held in entry 206 is added to the 32-bit offset 202 of the virtual address 200 by a summer 210 to yield a linear address 212.

The linear address 212 is then passed to a paging mechanism 214. The paging mechanism 214 includes page directories and tables that are used to translate the linear address 212 into a 32-bit physical address 216. The physical address 216 is used to specify location 218 within physical memory 106. In protected mode, since the physical address is 32 bits in length, up to 4 Gigabytes (i.e., 2^{32} Bytes) of memory can be addressed.

In V86 mode, an operating system may run real mode code in a virtualized manner while still providing some of the protection services found with protected mode. The operating system may provide a separate virtual machine for different real mode programs running on the microprocessor. Each virtual machine is an environment that is created through a combination of microprocessor capabilities and operating system software. Only one of the virtual machines is actively running on the microprocessor at any point in time. The operating system may context switch between virtual machines to change the currently active virtual machine. Every virtual machine must reside, ex-

ecute instructions, and refer to memory locations within the first megabyte of the virtual memory of the computer system.

The limitations of real mode and V86 mode pose a problem in a distributed environment. In order to appreciate the problem, it is helpful first to discuss what components play central roles in communications amongst computer systems in a distributed environment. Typically, each computer system in a distributed system has a network adapter for interfacing the computer system with the network. Figure 3 shows an example of a network adapter 314 that interfaces a computer 300 with a network 318. The computer 300 includes a central processing unit (CPU) 303 and a memory 301. The computer memory 301 contains a network driver 312 for the network adapter 314.

The network adapter 314 is responsible for sending packets to the network 318 and receiving packets from the network 318. A packet, in this context, is a packaged unit of information that is in a format proper for transmission over the network 318. The network adapter 314 includes a processor 315 and resident RAM 316 for helping it perform its duties.

The network driver 312 is responsible for transmitting packets that are destined for the network 318 and for receiving packets that are received from the network 318. The network driver 312 transmits and receives packets by buffering the packets. The problem posed by such buffering is that the network driver 312 is often a V86 mode driver and, as such, can only store the packets within the first megabyte of the virtual address space of computer memory 301.

Currently, the problem detailed above is addressed by providing a "Mapper" program. The Mapper program is situated between programs 306 that wish to receive packets or wish to send packets and the network driver 312. Programs 306 that wish to send or receive packets interface with the Mapper 308. The Mapper 308 has parts that execute in V86 mode and parts that execute in protected mode. The Mapper 308 provides buffers 310 for buffering the packets. In particular, the Mapper 308 allocates, frees and maintains linked lists of buffers 310 for interfacing between the programs 306 and the network driver 312. There is one list 302, 303, 304 for each size of buffers. In Figure 3, the buffers 302, 303 and 304 are sizes 1.5k, 256B and 64B, respectively. All buffers of the same size are linked on the same list.

Figure 4 shows a flow chart of the steps performed by the Mapper 308 during the process of receiving a packet from a program 306 and sending the packet to the network driver 312. The Mapper 308 allocates variable length buffers 310 within the first megabyte of memory (step 402 in

Figure 4) and stores them as linked lists (step 404). When a packet is received from the program (step 406), the Mapper 308 determines if there is a buffer large enough for the packet (step 408). If there is not a buffer large enough for the packet, the Mapper 308 waits until the network driver 312 has completed processing on a buffer that is large enough (step 410). If there is a buffer large enough for the packet, the Mapper 308 takes the smallest buffer that is large enough for the packet (i.e., the "best fit" buffer) off of a list (step 409) and stores the packet into the buffer (step 411). The size and number of the buffers is decided at the time of compilation of the Mapper 308. Next, the Mapper passes the buffer to the network driver 312 (step 412) and after the network driver 312 sends the packet and releases the buffer, the Mapper puts the buffer back into a linked list. The Mapper 308 ends processing upon a system shutdown (step 414). The Mapper 308, the network driver 312 and the buffers 310 all reside in global address space.

Problems arise when using the prior art method of buffering. As shown in Figure 3, when the Mapper 308 is performing a send, it receives a packet from a program 306 executing in protected mode, places the packet into a buffer 310 and passes the buffer to the network driver 312. The buffer 310 is not released by the network driver 312 and not able to be used by the Mapper 308 until the network driver 312 has completely finished performing the send function. Therefore, if the Mapper 308 has run out of buffers large enough for a packet, the Mapper 308 must wait idly until the network driver 312 frees up a buffer of sufficient size. When a buffer is finally available, execution must pass to protected mode to obtain the packet from the program 306, a copy must be performed into the buffer and then execution must pass back into V86 mode to pass the buffer to the network driver 312. This processing is very time consuming and the network adapter 314 could easily have completed sending all of the packets in its RAM 316 during this time, thus incurring idle time.

Problems also exist when packets are being received. As shown in Figure 3, if the Mapper 308 does not have a buffer 310 large enough for a receiving packet and must wait for the program 306 to release a buffer 310, the network adapter 314 may fill up its RAM 316 and experience an overflow of packets. When an overflow occurs, the overflowing packets are lost.

In addition to the problems noted above, using the "best fit" approach has another disadvantage. The Mapper 308, the buffers 310 and the network driver 312 all reside in global address space that is shared by all V86 tasks. When memory is used in the global address space, it cannot be used by other tasks. This makes global address space a

resource that should only be used sparingly. However, using the "best fit" approach leads to wasted memory since there will almost always be more space in the buffer than the exact size of the packet (i.e., there is internal fragmentation). Buffer sizes can be adjusted so as to minimize this waste, however, the anticipation of packet sizes is very difficult since different applications and network adapters use different packet sizes. Therefore, the "best fit" approach is not an optimal use of global address space.

Summary of the Invention

In accordance with a first aspect of the present invention, a method is executed in a distributed system having a number of computer systems that are interconnected by a network. In accordance with this method of the first aspect of the present invention, at least one of the computers contains a storage device, a network adapter and a V86 mode driver. During the transmission of packets, this method stores packets into a single buffer that resides within a portion of memory that is accessible by V86 mode drivers. Also during the transmission of packets, this method retrieves packets from the buffer while using a V86 mode driver.

In accordance with a second aspect of the present invention, the distributed system includes a network, a storage device and a processing means. The network provides for transmitting network packets. The storage device, which is connected to the network, has a portion that is accessible by V86 mode programs and contains a buffer. The processing means contains a packet loader for loading network packets into the buffer and a packet retriever for retrieving packets from the buffer.

Brief Description of the Drawings

Figure 1 depicts the addressing scheme employed for the 8088 microprocessor.

Figure 2 depicts the addressing scheme employed for the 80386 microprocessor executing in protected mode.

Figure 3 depicts a prior art system for buffering network packets.

Figure 4 depicts a flow chart of the steps performed in a prior art system of buffering network packets.

Figure 5 depicts a sample distributed system for practicing the preferred embodiment of the present invention.

Figure 6 shows a more detailed block diagram of a computer system of Figure 5.

Figure 7 depicts a circular buffer employed in the preferred embodiment of the present invention.

Figure 8 depicts a flow chart of the steps performed by the CopyDownPacket function in the preferred embodiment of the present invention.

Figure 9 depicts a flow chart of the steps performed by the FreeCircTxMem function in the preferred embodiment of the present invention.

Figure 10 depicts a flow chart of the steps performed by the ReceivePacket function in the preferred embodiment of the present invention.

Detailed Description of the Invention

The preferred embodiment of the present invention provides an improvement over the prior art for buffering packets when a V86 mode network driver is used. The preferred embodiment of the present invention improves memory utilization over the prior art by using a circular buffer that is accessible by the V86 mode network driver. Packet order is maintained within the circular buffer by the packets being processed in a first-in-first-out (FIFO) fashion. In addition, the preferred embodiment of the present invention provides higher system throughput by minimizing network adapter idle time and provides greater data reliability by minimizing the number of packets lost.

Figure 5 shows a distributed system 500 suitable for practicing the preferred embodiment of the present invention. The distributed system 500 includes computer systems 502, 504, 506, 508 and 510 that are interconnected by a network 511. The network 511 may be a local area network (LAN), a wide area network (WAN) or other suitable interconnection mechanism. Those skilled in the art will appreciate that the configuration for the distributed system 500 shown in Figure 5 is merely illustrative. The present invention may also be practiced on other distributed system configurations.

Figure 6 shows a more detailed block diagram of computer 502. Those skilled in the art will appreciate that the other computer systems 504, 506, 508, 510 may include similar components. The computer system 502 includes a central processing unit (CPU) 619, memory 600 and a network adapter 616. The network adapter 616 includes a processor 621 and resident RAM. The network adapter 616 generally takes the form of an expansion card and is connected to a network 511 that enables communications with other computers. The network adapter 616 is controlled by a network driver 614 that is resident in the memory 600 of the computer system 502. The network driver 614 is executed in V86 mode 604, thus it resides in the first megabyte of the virtual memory 600 of the computer system 502. Programs 606 act as originators and end recipients of packets that are transmitted across the network 511. The programs 606 may be protocol drivers. The programs 606 generally execute

in protected mode 602 and, thus, do not have to reside within the first megabyte of memory. The network driver 614 and the programs 606 exchange packets through a Mapper 608. The Mapper 608 includes portions that execute in protected mode and portions that execute in V86 mode. Buffers 613 are provided in the first megabyte of memory to buffer packets between the programs 606 and the network driver 614. The buffers 613 include a circular send buffer 612 and a circular receive buffer 610 that will be described in more detail below.

Although the programs 606 are described as being the originators and the end recipients of packets, one skilled in the art will recognize that the programs may be code segments, processes, libraries or other entities that are capable of generating packets or receiving packets. In addition, although the description of the present invention refers to the use of packets, one skilled in the art would recognize that the present invention may also be used with other forms of data. Further, one skilled in the art would recognize that the circular buffer could be implemented so that the size of the circular buffer changes at runtime.

Although Figure 6 shows only a single circular send buffer 612 and a single circular receive buffer 614, it should be appreciated that larger numbers of buffers may be employed. For example, when the network driver 614 is an Open Data Link Interface (ODI) network driver, distributed by Novell Corporation of Provo, Utah, one circular send buffer 612 and one circular receive buffer 610 are required. Further, when the network driver 614 is a Network Driver Interface Specification (NDIS) V86 mode driver, distributed by Microsoft Corporation of Redmond, Washington, one circular send buffer 612 is used for all network drivers; however, one non-circular buffer is used as a receive buffer per adapter. The receive buffer is not circular because the NDIS network driver provides synchronous communications for packet reception and never receives more than one packet at a time.

The circular send buffer 612 and the circular receive buffer 610 are of like configuration. Figure 7 shows the configuration of such buffers. These circular buffers 610 and 612 are pre-allocated in the first Megabyte of memory at the time of system start-up and are maintained during the life time of the system by the Mapper 608. Each of the circular buffers 710 is of a fixed length. In Figure 7, the buffer 710 has a length of 4k (i.e., 4096 bytes). The buffer 710 has a physical beginning and a physical end as well as a logical beginning ("head") and a logical end ("tail"). Pointers are maintained to specify the logical beginning and the logical end of the buffer 710. The pointer TxStart 700 points to the logical head of the circular buffer 710, whereas the pointer TxStop 702 points to the

logical tail of the circular buffer 710. Other status information is maintained for the circular buffer 710. This status information includes a value designated as TxLen, which equals the total length of the buffer 710; a value TxAvail, which specifies the amount of buffer space available; and a value known as Requests, which specifies the number of packets currently in the buffer 710.

When a packet is stored in a circular buffer 710, a header is added to the packet. The header includes two fields 704 and 706, as shown in Figure 7. The first two bytes of the header contain the In Use ("IU") flag 704 which signifies whether this portion of the buffer 710 is still in use or is free to be used by another packet. The second field, the Len (short for length) field 706, is also two bytes long and specifies the length of the data 708 in this portion of the buffer 710. The data field 708 is variable length depending on the size of the packet.

The Mapper 608 is responsible for writing packets received from the programs 606 into the circular buffers 610. The Mapper 608 calls a CopyDownPacket routine in order to copy a packet from the programs 606 into a buffer that is then submitted to the network driver for transmission across the network. By using the CopyDownPacket routine, memory utilization is enhanced because there is no internal fragmentation within the buffer. In addition, system throughput is increased due to the Mapper 608 not being limited to a fixed number of fixed sized buffers, thus the only time that the Mapper must wait for a portion of the buffer to become available is when the buffer is entirely full, in which case throughput does not suffer. Figure 8 shows a flow chart of the steps performed by the CopyDownPacket Routine. The CopyDownPacket routine first determines if there is sufficient room in the buffer 710 to store the packet (step 800). This determination is done by comparing the packet size with the value TxAvail, which specifies the available space in the buffer 710. If there is not enough available space in the buffer 710, then the Mapper 608 waits and does not send the request ("sends the request") (step 802) and the request is tried again later. If there is sufficient room in the buffer 710, then the In Use flag 704 is set to indicate that the portion of the buffer to be occupied by the packet is in use (step 804). In addition, the amount of memory in the buffer 710 that the packet occupies is written into the len field 706. Next, TxStop 702 is incremented by the size of the packet plus four bytes, to account for the header information (step 808). If when TxStop 702 is incremented, the value of TxStop 702 exceeds the buffer range (i.e., beyond the physical end of the buffer) (see step 810), the value of TxStop is set to wrap around to the beginning of the buffer

(step 812). Next, the packet is copied into the data field 708 of the allocated portion of buffer 710 (step 814). Finally, the value TxAvail is decremented by the amount of buffer space occupied by the packet to reflect the storage of the packet therein (step 816). The "Requests" variable is incremented by one to indicate that another request has been processed (step 818), and the Mapper 608 passes the address of the portion of the buffer to the network driver 614 (step 820). After the network driver 614 receives the address of the portion of the buffer, the network driver 614 copies the packet out of the buffer and then copies the packet into the RAM of the network adapter 616. The network adapter 616, in turn, sends the packet across the network 511.

After a packet has been sent from the circular receive buffer 610 or the circular send buffer 612, the Mapper needs to free up the buffer space that held the packet so that the space may again be used. The FreeCircTxMem routine performs this task. Figure 9 shows a flow chart of the steps performed by the FreeCircTxMem Routine. First, the "In Use" flag 704 is set to "not in use" for the packet that has been sent (step 900). Then, if the TxStart pointer 700 is pointing to the beginning of the buffer space that held the packet (see step 902), the TxStart pointer 700 is incremented by the len field 706 plus four (step 904) and TxAvail is incremented by the len field 706 plus four (step 905). If TxStart 700 previously pointed to the space that the packet resided in and if, after being incremented, the new TxStart 700 points to a portion of the buffer that is not in use (step 906), TxStart 700 is again incremented by the value in the len field 706 (step 904) and TxAvail is again incremented by the value in the len field 706 plus four (step 905). The incrementing of TxStart 700 continues until it is determined that the buffer 710 is entirely empty or until a portion of the buffer 710 is found that has the in use field 704 marked as "In Use." Next, the "Requests" variable is decremented by one (step 910). However, if the TxStart pointer 700 does not point to the portion of the buffer 710 that the packet resided in (step 902), the Mapper merely decrements the "Requests" variable (step 910).

The circular receive buffer 610 must also be able to receive packets. The circular receive buffer 610 receives packets from the network 511 by using the ReceivePacket routine. The ReceivePacket routine accepts a packet from the network driver 614 and passes it to the programs 606. When an asynchronous network driver is receiving a packet, such as the ODI network driver, the driver sends the Mapper 608 a "ReceiveLookAhead" message. This message indicates that a packet is being received, and the message also contains up to the first 128 bytes of the packet. After the network driver 614 has received the entire packet,

the driver sends a "Complete" message to the Mapper 608 to indicate that the entire packet has been received. A complication arises because the communications through the network adapter 616 and network driver 614 are occurring asynchronously, hence, it is possible that "complete" messages may arrive out of the order of the "ReceiveLookAhead" messages. Therefore, if every time a complete is received by the Mapper 608, the Mapper 608 hands off the packet to the programs 606, the packets will fall out of order. The present invention overcomes this problem.

Figure 10 shows a flow chart of the steps performed by the ReceivePacket routine. When the ReceivePacket routine first receives a ReceiveLookAhead packet from the network driver 614, the Mapper 608 determines if there are any packets for which "Complete" messages have not been received (step 1000). If there are no "Complete" messages outstanding, then the Mapper determines if the packet is less than 128 bytes (step 1007). If the packet is less than 128 bytes, the packet is sent directly to the program 606 (step 1003), no buffering is needed. If the packet is greater than 128 bytes (step 1001), the Mapper 608 allocates a portion of the buffer, passes the address of the portion of the buffer to the network driver 614 (step 1002), waits for the network driver 614 to complete filling in the packet into the portion of the buffer (step 1004), and then sends the address of the portion of the buffer to the programs 606 (step 1018). However, if there are complete messages that are outstanding, when the Mapper 608 receives a ReceiveLookAhead message, the Mapper 608 marks the next available portion of the circular buffer 710 in the In Use field 704 as a quick buffer (step 1006). Marking a portion of the buffer as a quick buffer signifies that the address of the buffer cannot be sent to the programs 606 until all of the preceding outstanding complete messages have been received. By using a quick buffer, the Mapper 608 ensures that the order of the packets is maintained, thus providing data integrity. Next, the Mapper 608 determines if the packet is less than 128 bytes (step 1008). If the packet is less than 128 bytes, the Mapper 608 automatically copies the packet into the quick buffer (step 1010). Otherwise, the Mapper 608 must create an ECB 1012 (memory block descriptor), which contains an address to a quick buffer, and passes the ECB to the network driver 614. In addition, the Mapper 608 will store the address of the ECB in the "Len" field 706 of the quick buffer (step 1012). The network driver 614 next fills in the portion of the buffer with the received packet and notifies the Mapper 608 when it has completed (step 1014). Upon receiving the "Complete" message for this receive, the Mapper 608 must wait for all of the previous outstand-

ing complete messages to be received (step 1016) before it can pass the address of the portion of the buffer up to the programs 606 (step 1018). Finally, when all outstanding complete messages have been received, the Mapper 608 then passes the address of the quick buffer wherein the packet resides to the programs 606 (step 1018).

While the present invention has been described with reference to a preferred embodiment thereof, those skilled in the art will know of various changes in form that may be made without departing from the spirit and scope of the claimed invention as defined in the appended claims.

Claims

1. In a distributed system having computer systems, at least one of said computer systems having a storage device, a network adapter and a V86 mode driver for interfacing with the network adapter, said distributed system including a network for transmitting network packets, a method comprising the steps of:
 - storing network packets in a single buffer in a portion of the memory that is accessible by the V86 mode driver; and
 - retrieving network packets from the buffer using the V86 mode driver.
2. A method as recited in claim 1 wherein the buffer is comprised of a sequence of portions and has a beginning and an end, wherein the step of storing network packets includes the steps of:
 - while storing a packet,
 - storing information contained in the packet in sequential portions in the buffer;
 - detecting if the end of the buffer is reached before all of the information contained in the packet is stored; and
 - if the end of the buffer is reached before all of the information in the packet is stored, storing the information contained in the packet that has not yet been stored in the buffer in sequential positions starting at the beginning of the buffer.
3. A method as recited in claim 1 wherein the step of storing network packets includes the step of storing each packet in a next sequential section of the buffer based upon the order in which the packet was received at the buffer.
4. A method as recited in claim 1 wherein said V86 mode network driver interfaces with a protected mode computer program, and the method further comprising the step of the buffer receiving said network packets from the protected mode computer program.
5. A method as recited in claim 1, further comprising the step of the buffer receiving said network packets from the V86 mode network driver.
6. A method as recited in claim 1 wherein said V86 mode network driver interfaces with a protected mode computer program, and the method further comprising the step of sending said network packets from the buffer to the protected mode computer program.
7. A method as recited in claim 1, further comprising the step of sending said network packets from the buffer to the V86 mode network driver.
8. In a distributed system, the distributed system comprising of a plurality of computer systems, each computer system being interconnected, at least one computer system having a memory and a microprocessor, the microprocessor having the ability to run in V86 mode and protected mode, a method for sending and receiving packets comprising the steps of:
 - allocating at least one fixed-length buffer in a portion of the memory that is accessible in V86 mode;
 - storing a plurality of packets in the buffer, each packet being stored sequentially; and
 - retrieving packets from the buffer.
9. A method as recited in claim 8 wherein the buffer is comprised of a sequence of portions and has a beginning and an end, wherein the step of storing a packet includes the steps of:
 - while storing a packet,
 - storing information contained in the packet in sequential portions in the buffer;
 - detecting if the end of the buffer is reached before all of the information contained in the packet is stored; and
 - if the end of the buffer is reached before all of the information in the packet is stored, storing the information contained in the packet that has not yet been stored in the buffer in sequential positions starting at the beginning of the buffer.
10. A method as recited in claim 8 wherein the buffer comprises a sequence of sections and wherein the step of storing network packets includes the step of storing each packet in a next sequential section of the buffer based upon the order in which the packet was received at the buffer.

11. A method as recited in claim 8 wherein the buffer has a logical tail, wherein the method further comprises the steps of maintaining a count of the number of packets in the buffer, and maintaining a count of available space in the buffer; wherein the step of storing packets includes the steps of updating the pointer to the logical tail of the buffer, incrementing the count of the number of packets in the buffer and incrementing the count of available space within the buffer.
12. A method as recited in claim 8 wherein the buffer has a logical head wherein the method further comprises the steps of maintaining a count of the number of packets in the buffer; and maintaining a count of available space within the buffer; wherein the step of retrieving packets includes the steps of updating a pointer to the logical head of the buffer, decrementing the count of the number of packets in the buffer and decrementing a count of available space within the buffer.
13. A distributed system comprising:
 a network for transmitting network packets;
 a storage device connected to the network having a portion that is accessible by V86 mode programs and said portion holding a buffer for buffering network packets; and
 a processing means connected to the network comprising:
 a packet loader for loading network packets into the buffer; and
 a packet retriever for retrieving packets from the buffer.
14. A computer system for sending and receiving packets, comprising:
 a network connected to the computer system;
 a network adapter connected to the network for sending and receiving packets across the network;
 a network driver for sending and receiving packets from the network adapter;
 a computer program for sending packets to the network driver and for receiving packets from the network driver, said computer program operating in protected mode;
 a circular buffer for holding packets that are being received from the computer program and for holding packets that are destined for the network driver, said circular buffer being accessible by programs executing in V86 mode; and
 a circular buffer for holding packets that are being received from the network driver and for holding packets that are destined for the computer program, said circular buffer being accessible by programs executing in V86 mode.
15. A method in a computer system for sending packets, the computer system having a memory and a microprocessor, the microprocessor having the ability to execute in V86 mode and protected mode, the computer having a circular buffer, the circular buffer having a physical beginning and a logical end, the packets consisting of a plurality of bytes, comprising the steps of:
 while the microprocessor is executing in V86 mode, allocating a fixed-length circular buffer in the memory of the computer; and
 for each packet,
 determining if the circular buffer has enough available space for the packet;
 storing an indication that signifies whether a portion of the buffer is being used or not at the next location after the logical end of the circular buffer;
 storing the length of the packet after the indication;
 for each byte in the packet,
 if sufficient free buffer space exists for the byte after the logical end, copying the byte to the next location after the logical end of the circular buffer;
 if there is not sufficient free buffer space after the logical end of the buffer and before the physical end of the buffer for the byte, copying the byte to the physical beginning of the circular buffer; and
 updating the logical end of the buffer.
16. A method in a computer system for receiving packets, the computer system having a memory and a microprocessor, the microprocessor having the ability to run in V86 mode and protected mode, the computer system having a circular buffer and a network driver, the circular buffer having a physical beginning and a logical end, the circular buffer interfacing with the network driver, the network driver sending a receive complete message when a packet has been received, the packets consisting of a plurality of bytes, the method comprising the steps of:
 while the microprocessor is executing in V86 mode,
 allocating a fixed-length circular buffer in the memory of the computer; and
 for each packet,
 determining if the circular buffer has enough available space for the packet;

storing an indication that signifies whether a portion of the buffer is being used or not at the next location after the logical end of the circular buffer;

storing the length of the packet after the indication 5

for each byte in the packet,

if sufficient free buffer space exists for the byte after the logical end, copying the byte to the next location after the logical end of the circular buffer; 10

if there is not sufficient free buffer space after the logical end of the buffer and before the physical end of the buffer for the byte, copying the byte to the physical beginning of the circular buffer; and 15

updating the logical end of the buffer.

17. A method as recited in claim 16 wherein the step of storing the in use flag includes the step of detecting if a receive complete is outstanding and storing an identifier signifying that the circular buffer is a quick circular buffer. 20

18. A method in a computer system for freeing memory associated with a packet contained in a circular buffer, the computer system having a microprocessor, wherein the microprocessor has the ability to run in V86 mode and protected mode, the circular buffer contains a plurality of packets, each packet resides in a different portion of the circular buffer, each portion of the circular buffer has an in use flag that may be set as in use or not in use, and a length field, the circular buffer has a logical head, a logical tail, a count of the available memory in the circular buffer and a count of the number of packets in the circular buffer, the method comprising the steps of: 25

while the microprocessor is executing in V86 mode, 40

for each portion of the circular buffer to be freed,

setting the in use flag to not in use;

if the logical head of the circular buffer points to the portion of the buffer, for each subsequent portion of the buffer, 45

if the in use flag of the subsequent portion is set to not in use, increasing the logical head pointer by the value in the length field of the subsequent portion; 50

decreasing the count of the number of packets by the number of the portions of the buffer that were freed; and

decreasing the count of the available memory by the amount of memory that was allocated for all of the portions of the buffer that were freed. 55

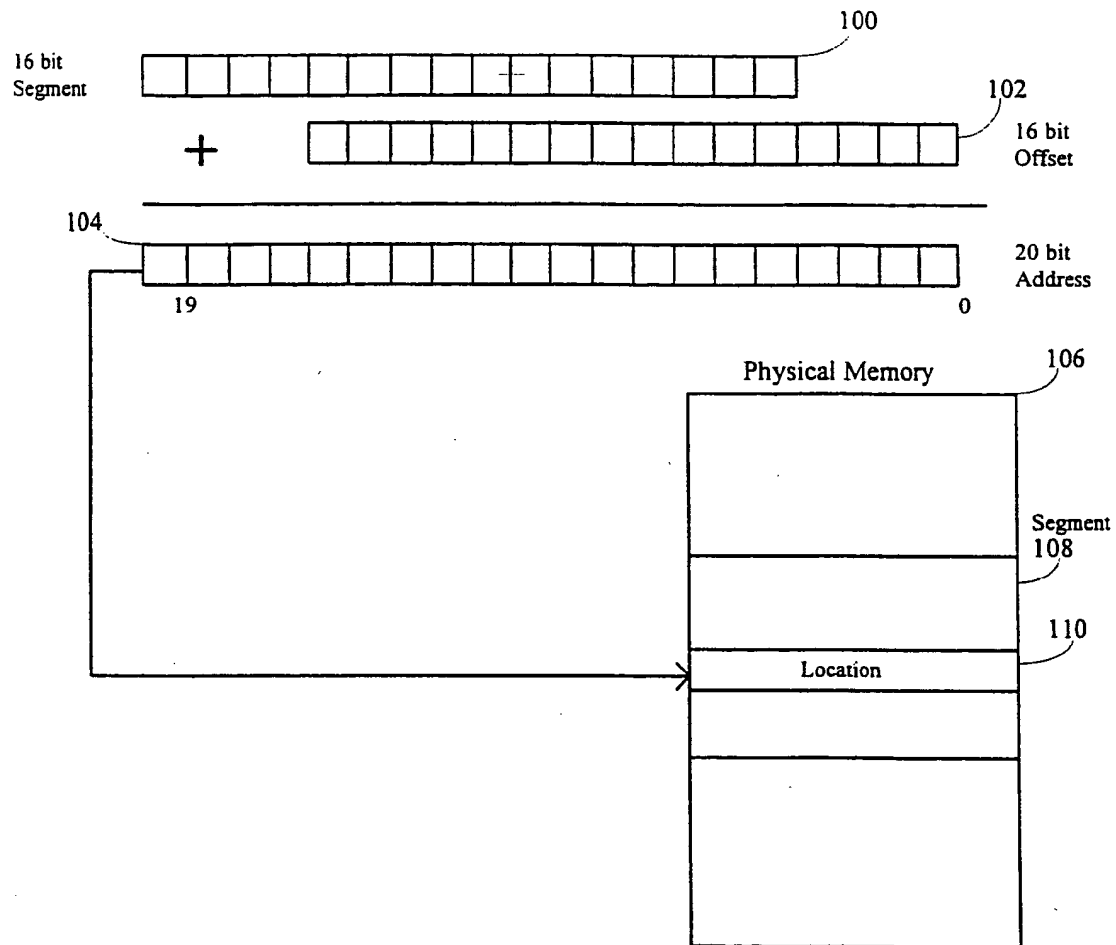


Figure 1
(Prior Art)

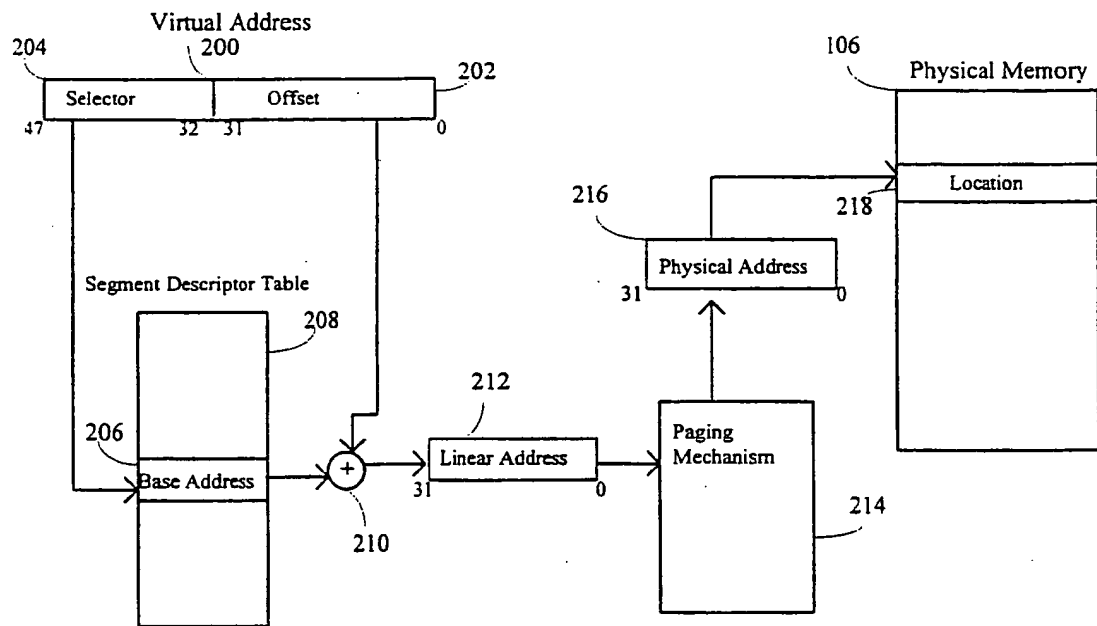


Figure 2
(Prior Art)

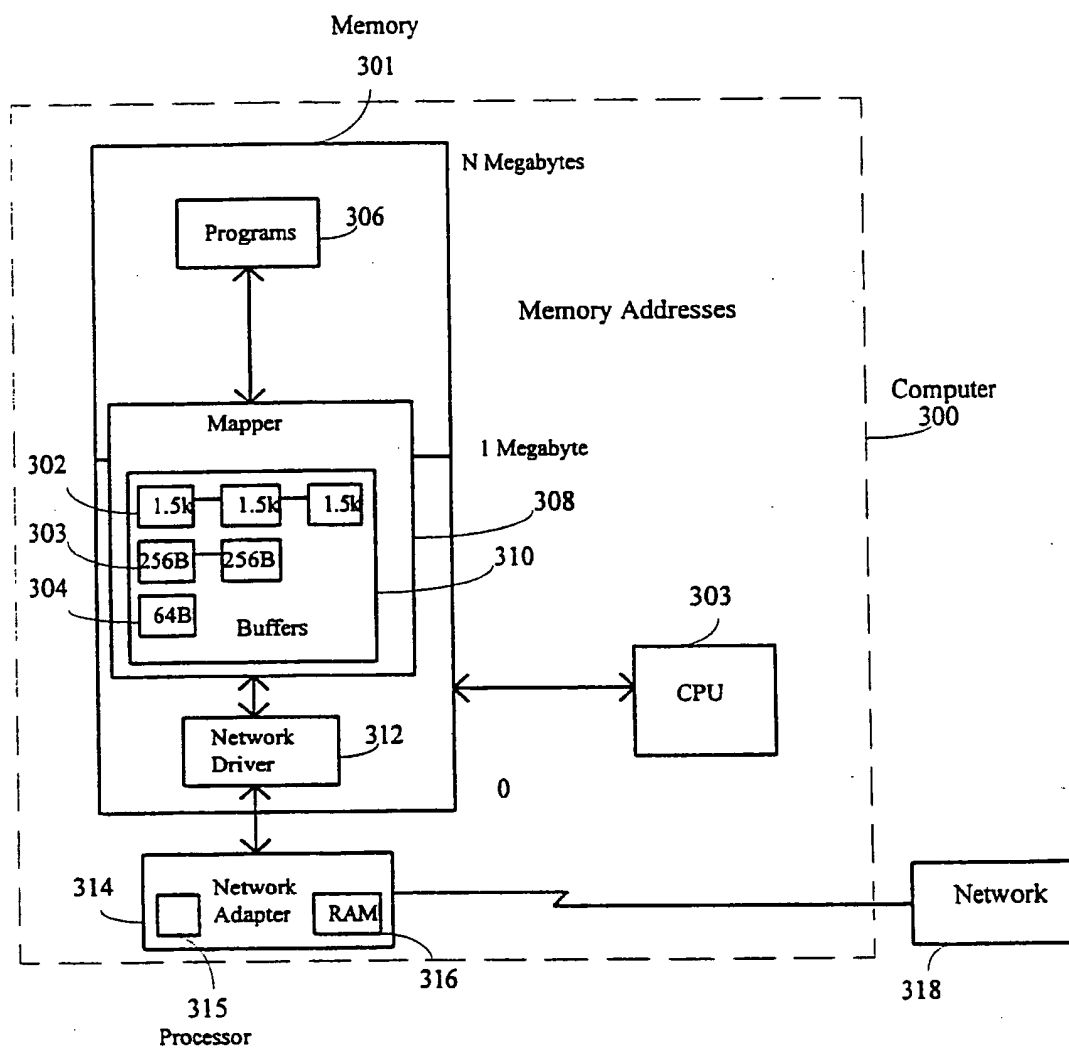


Figure 3
(Prior Art)

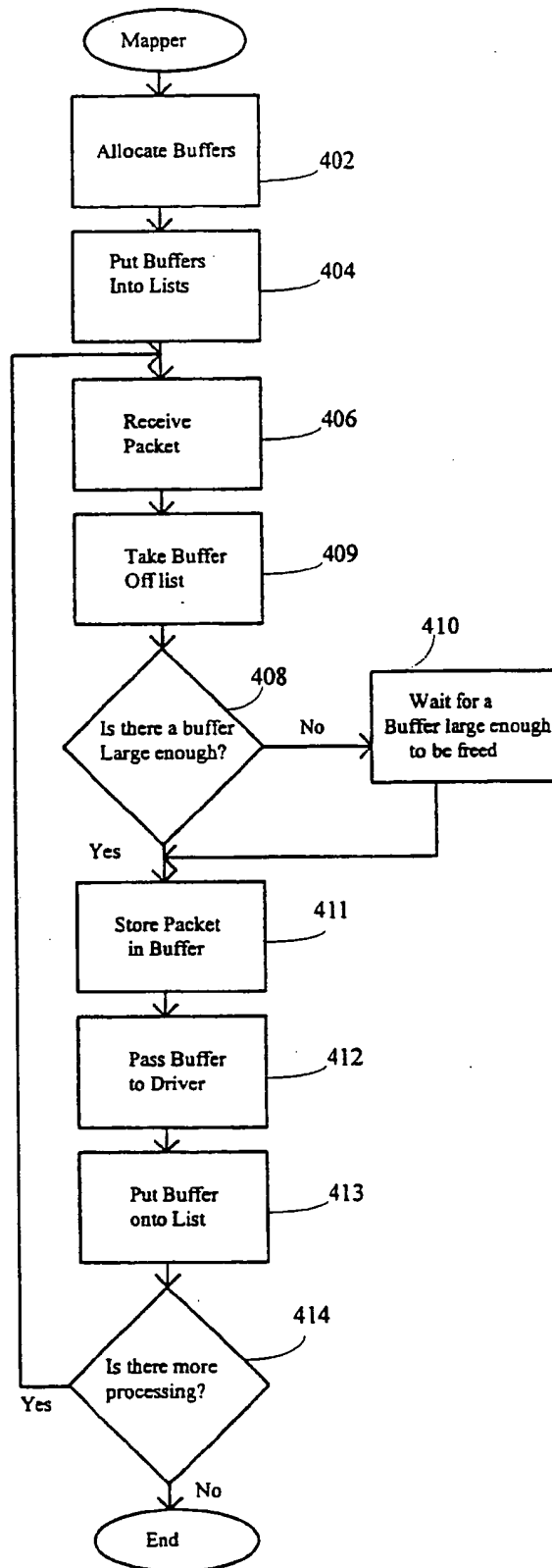


Figure 4
(Prior Art)

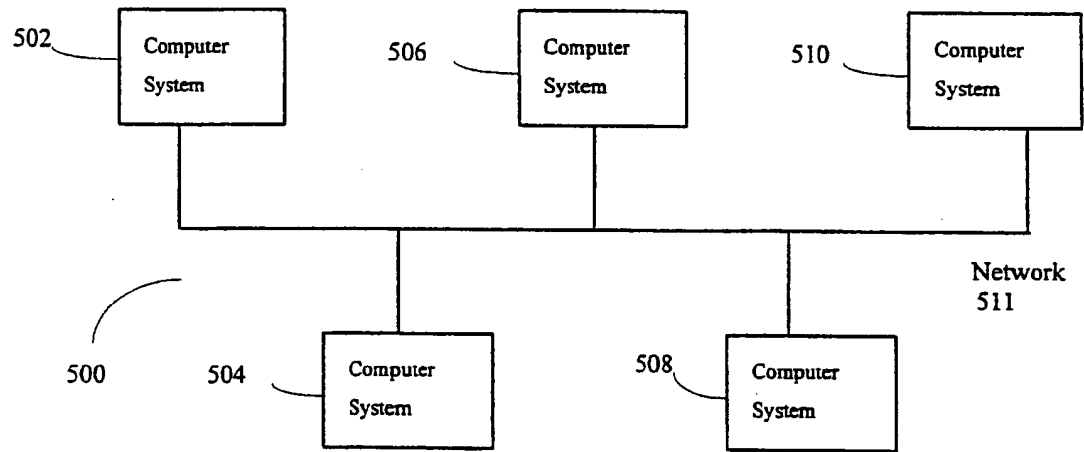


Figure 5

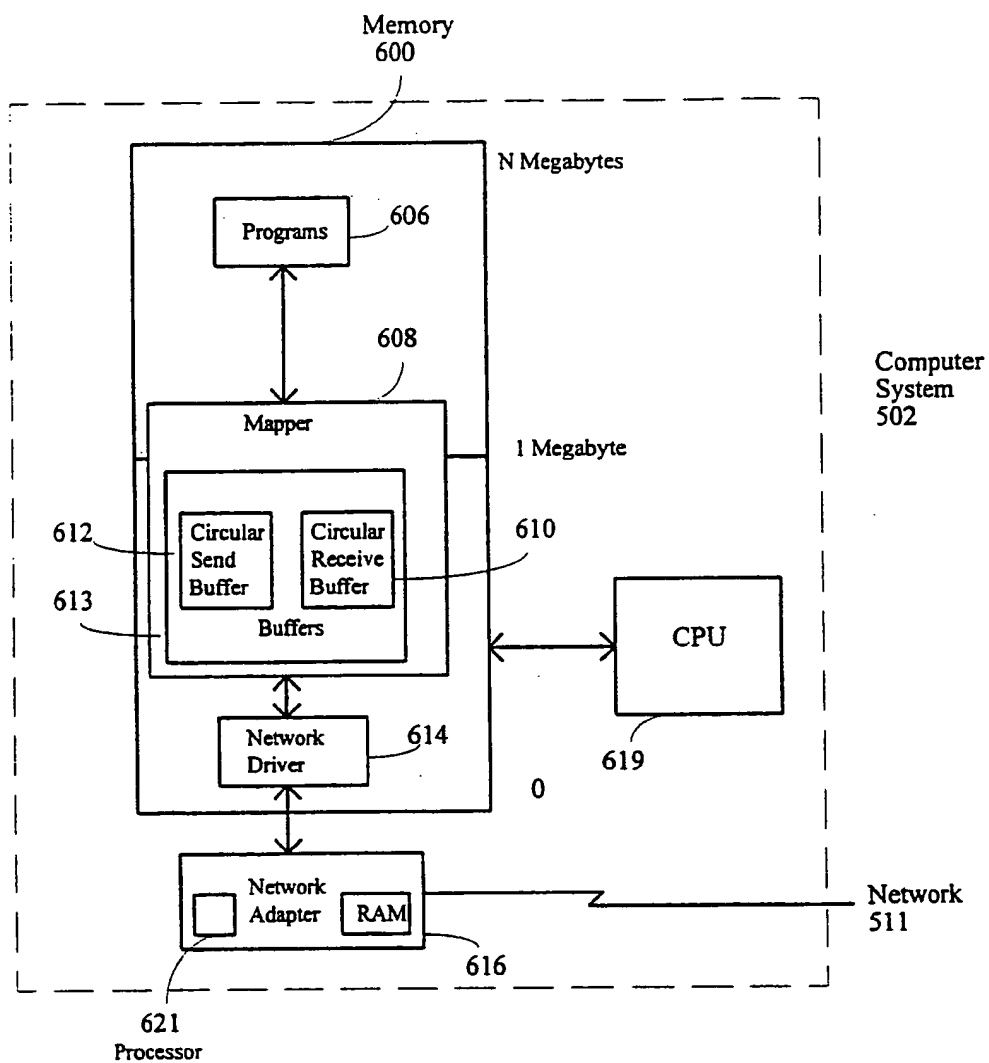


Figure 6

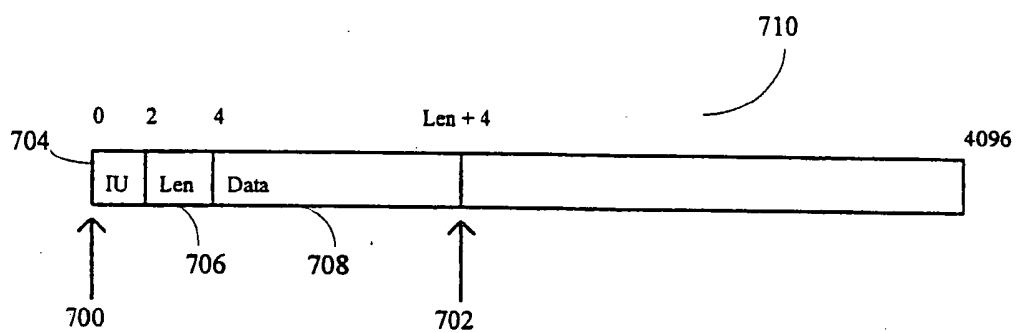


Figure 7

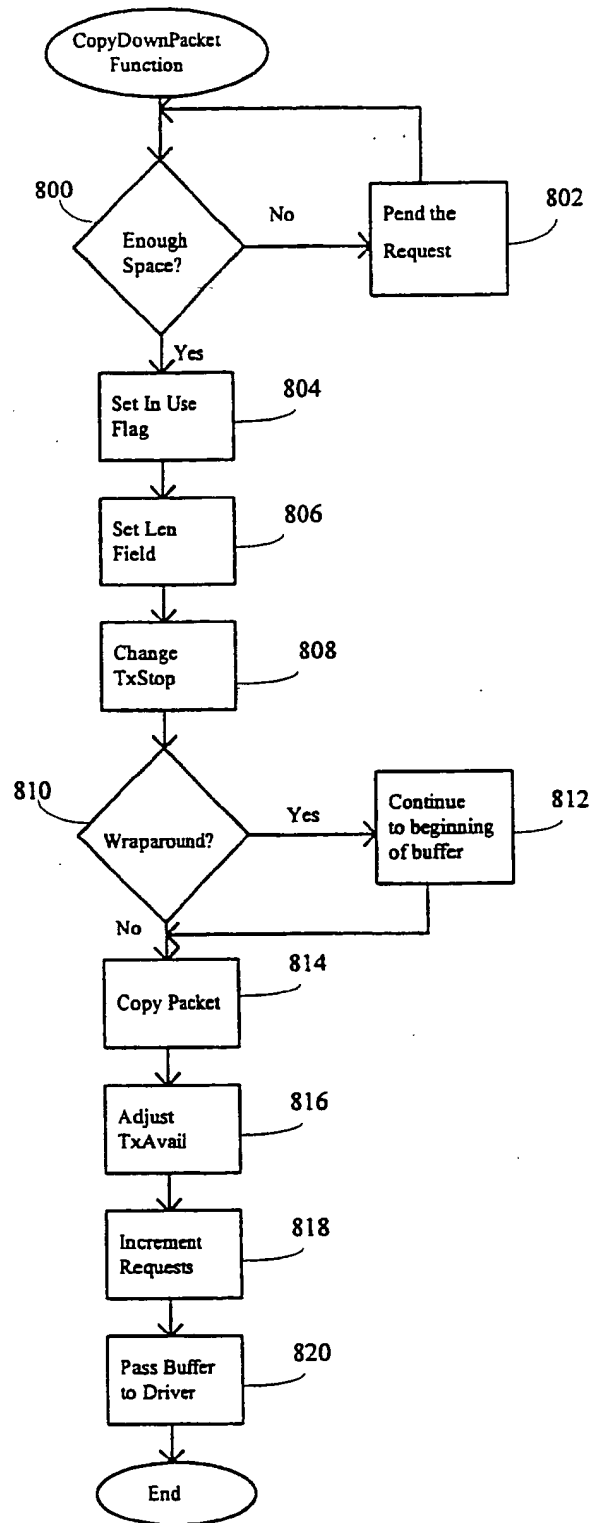


Figure 8

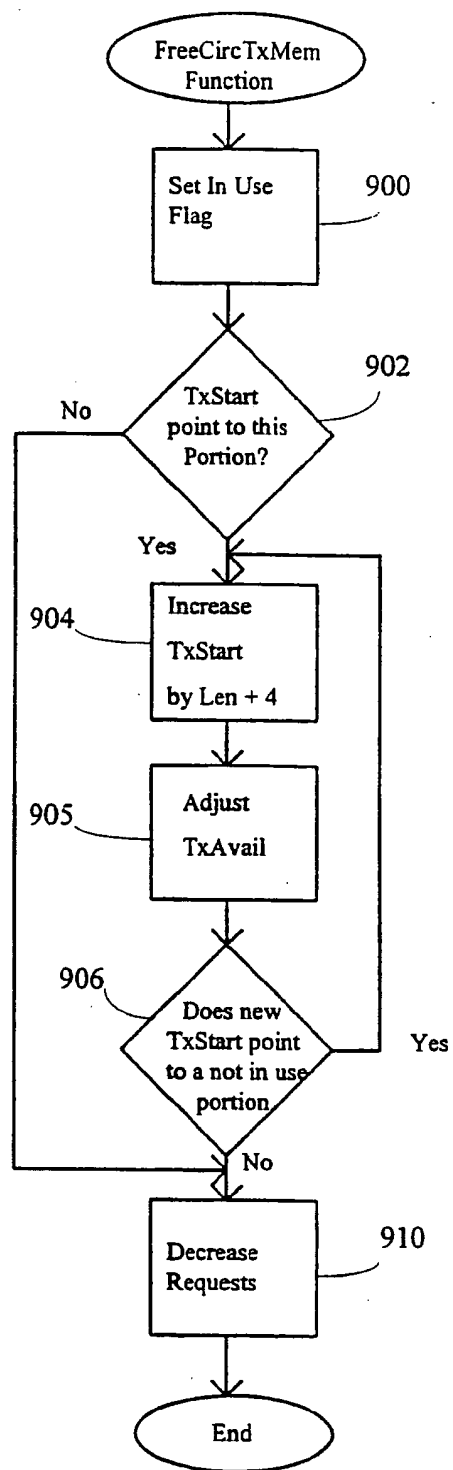


Figure 9

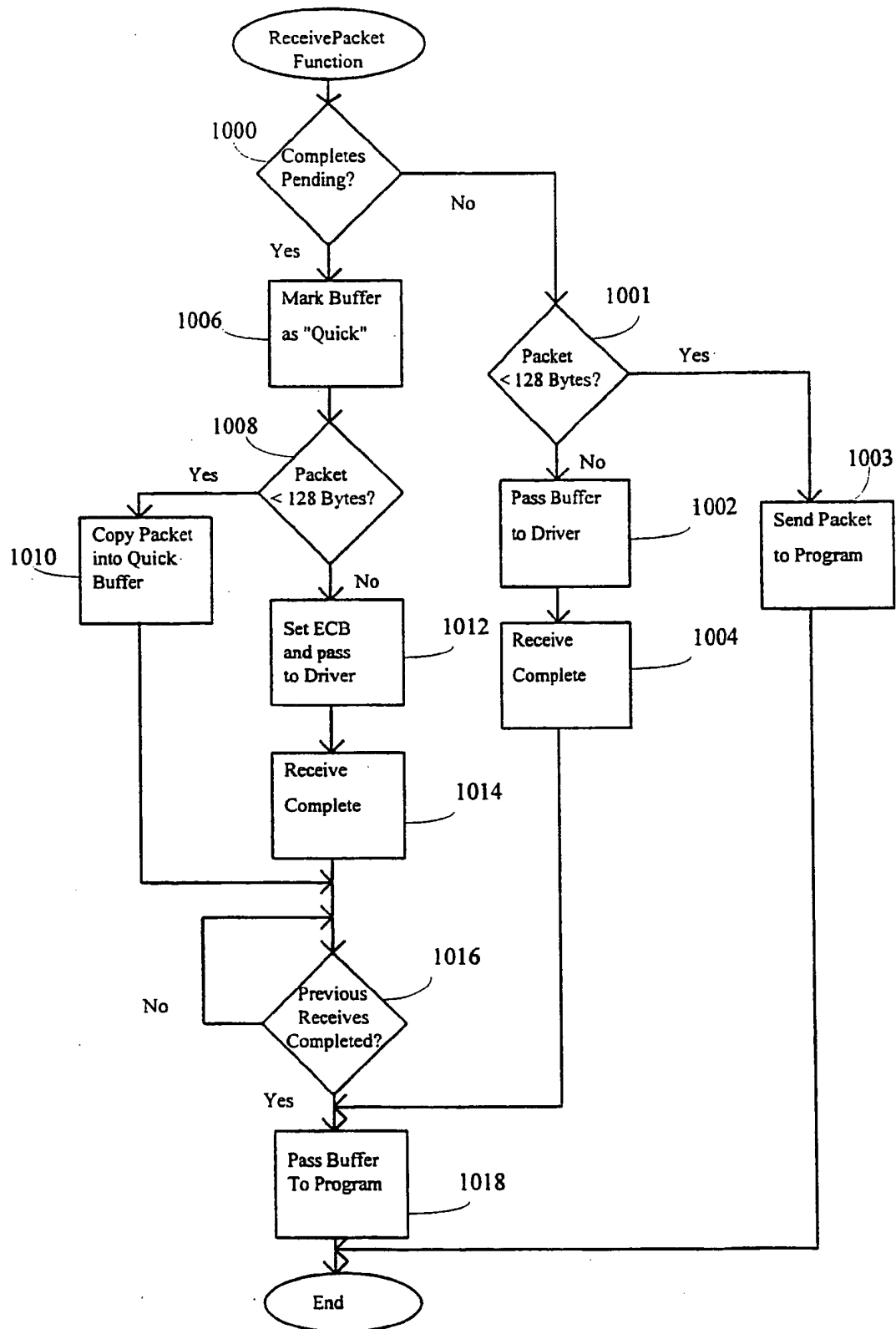


Figure 10

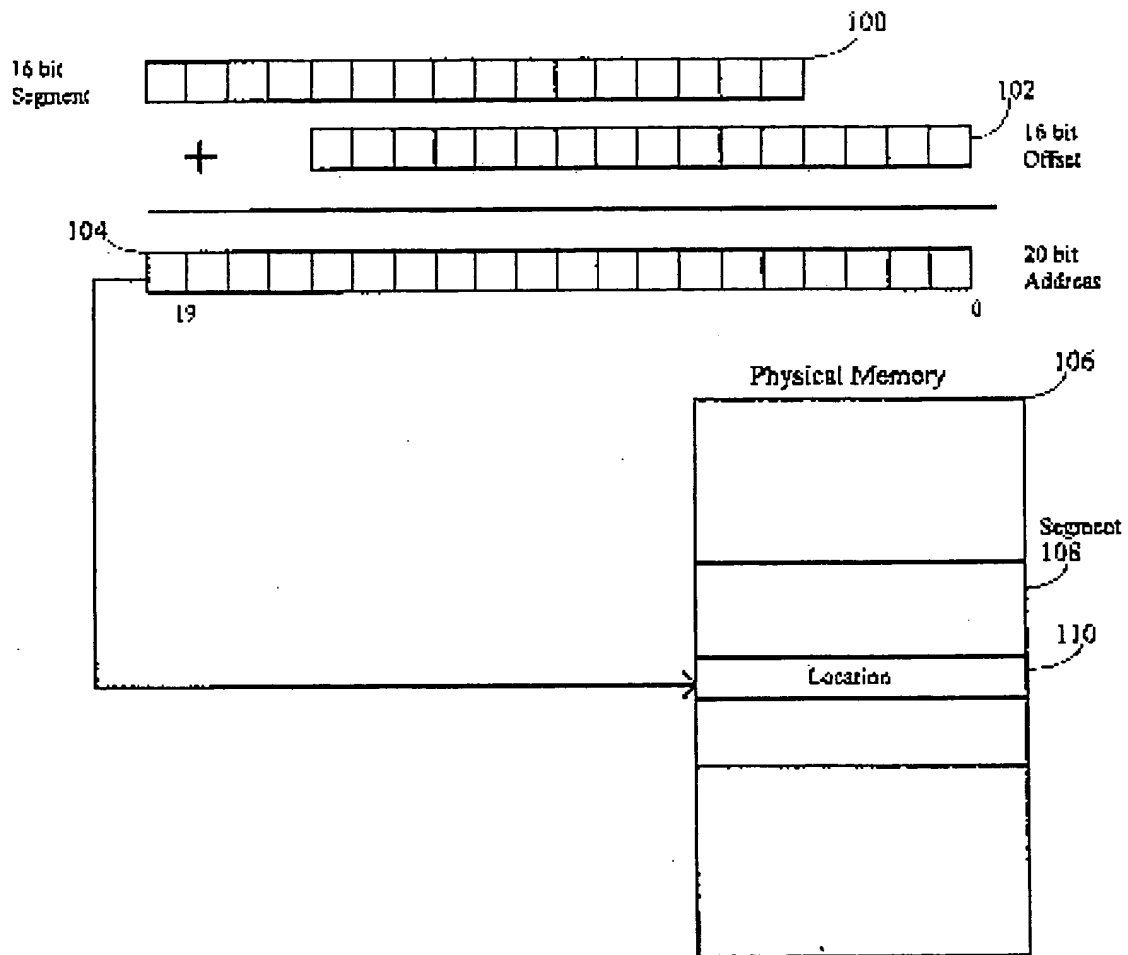


Figure 1
(Prior Art)

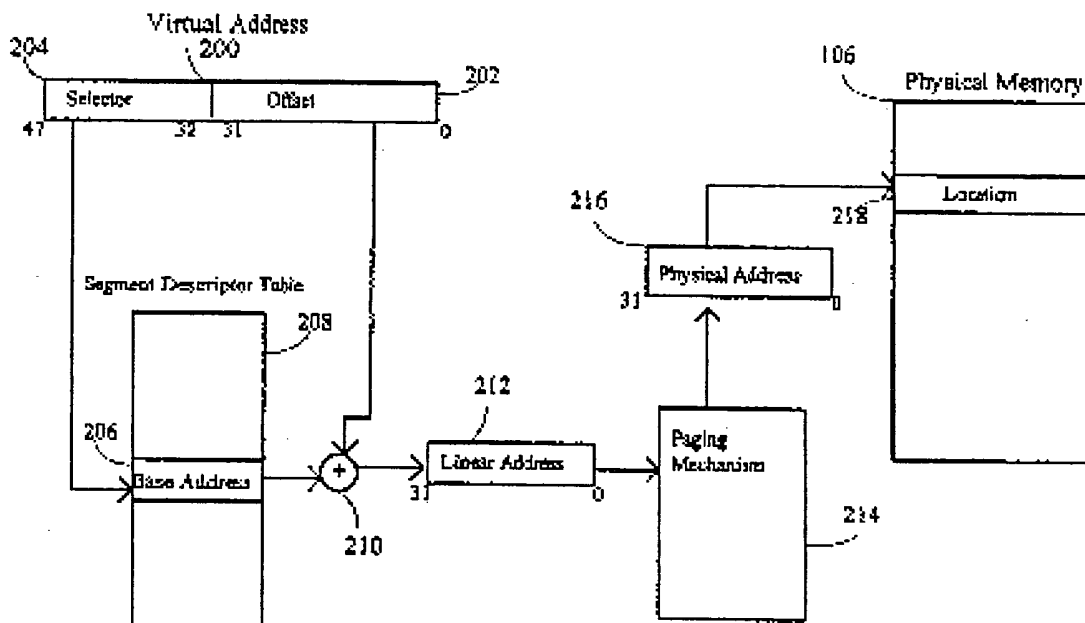


Figure 2
(Prior Art)

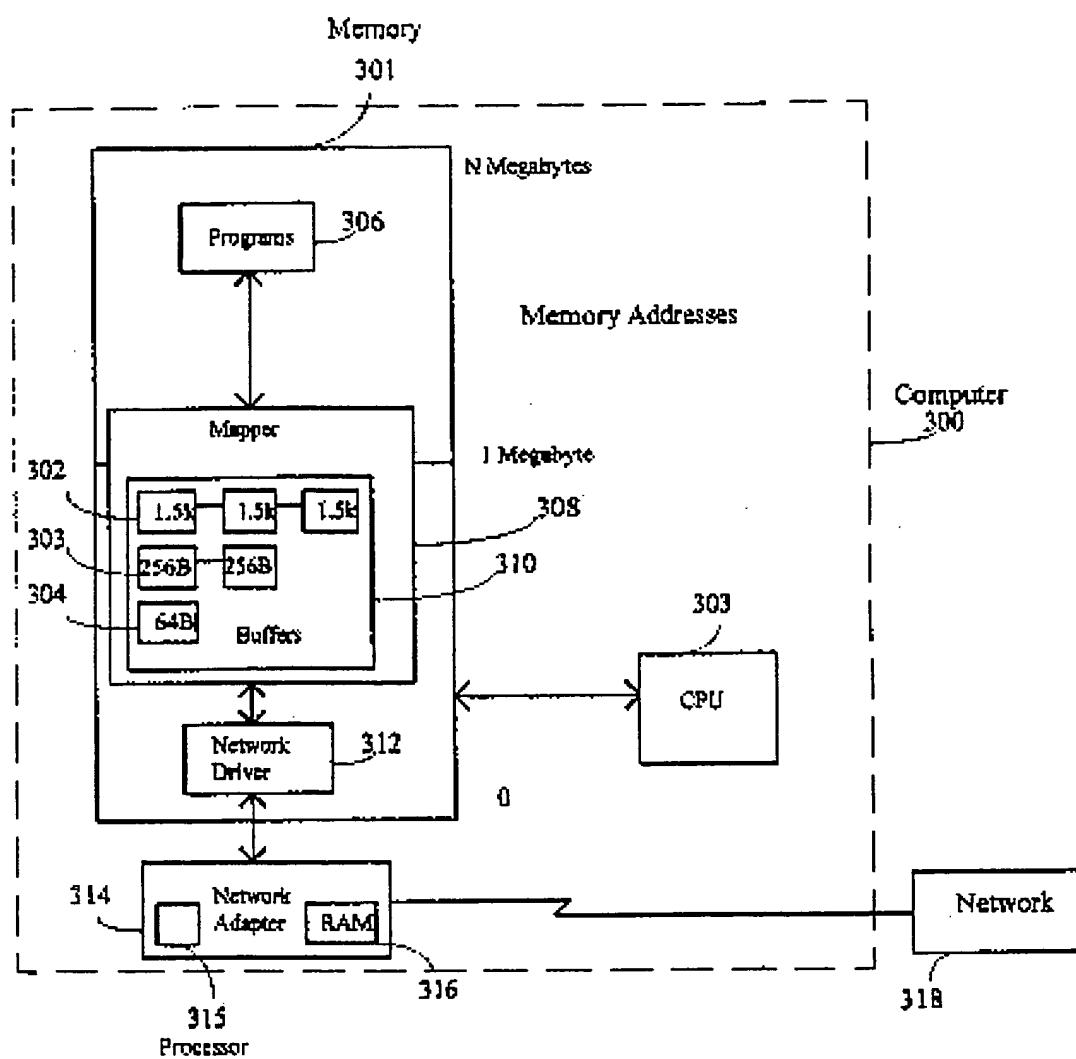


Figure 3
(Prior Art)

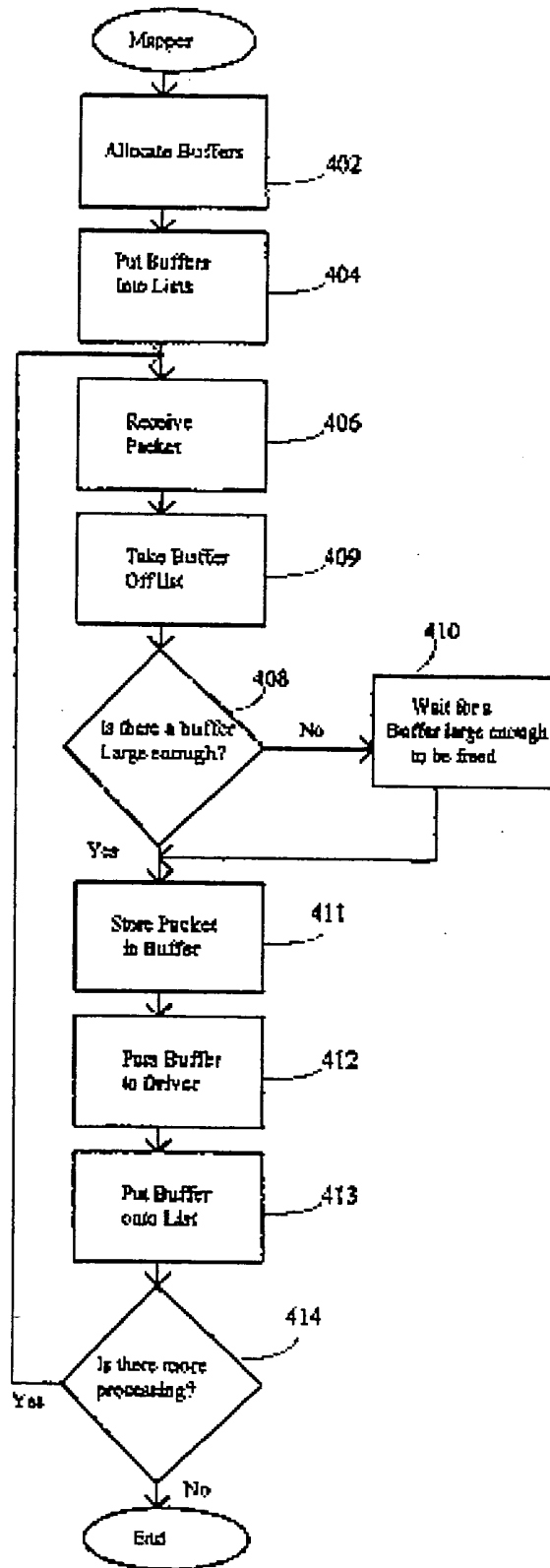


Figure 4
(Prior Art)

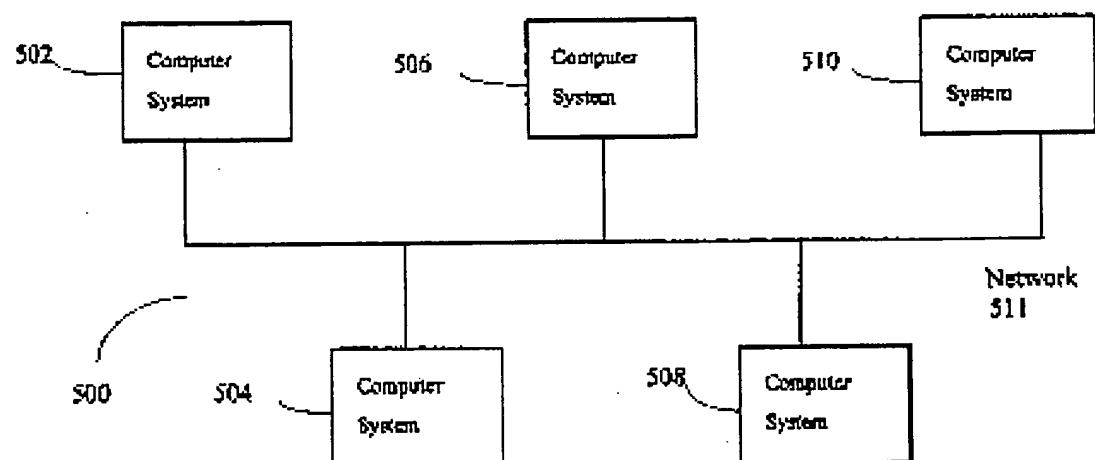


Figure 5

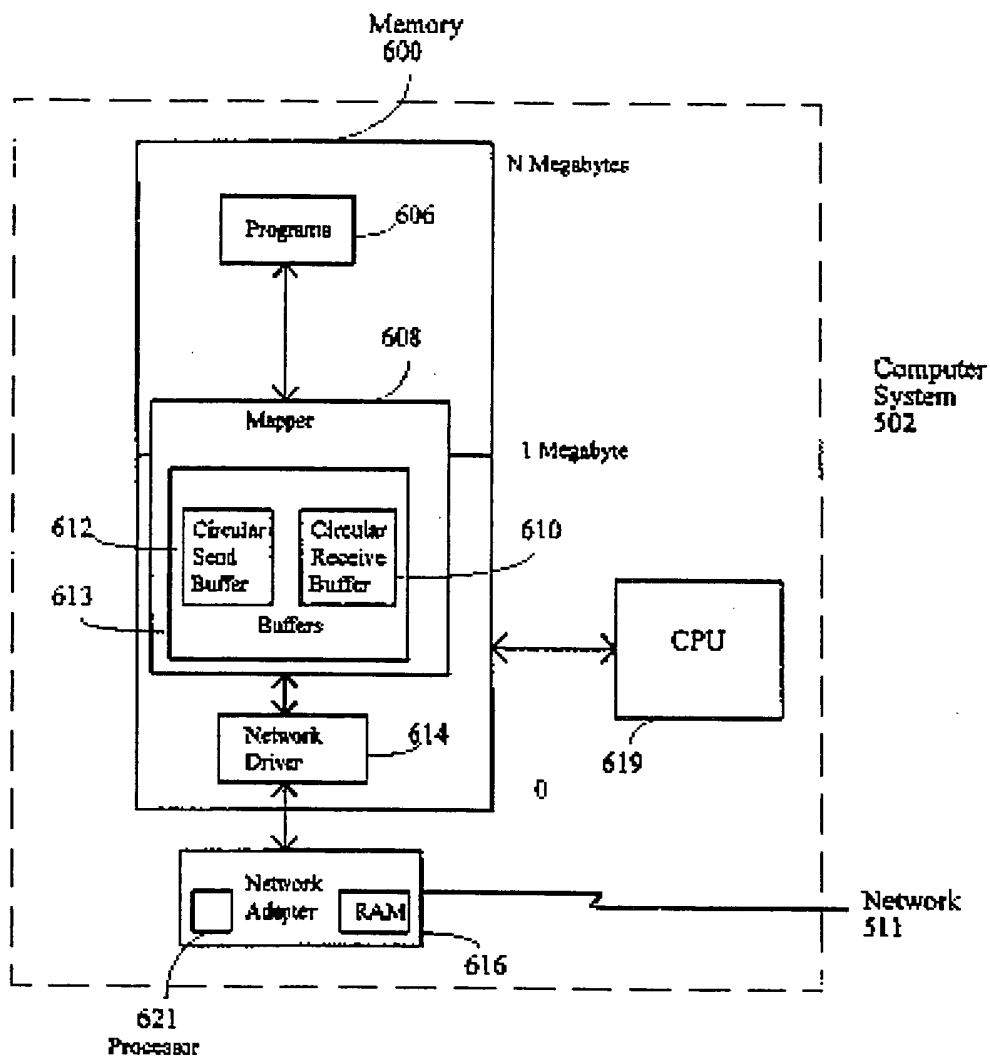


Figure 6

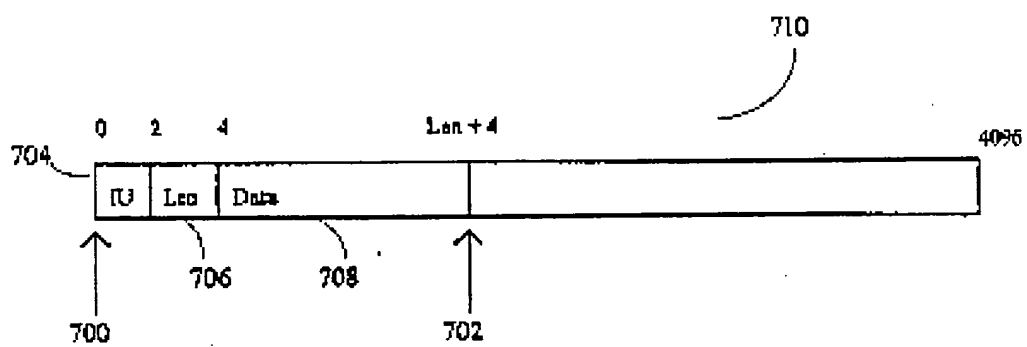


Figure 7

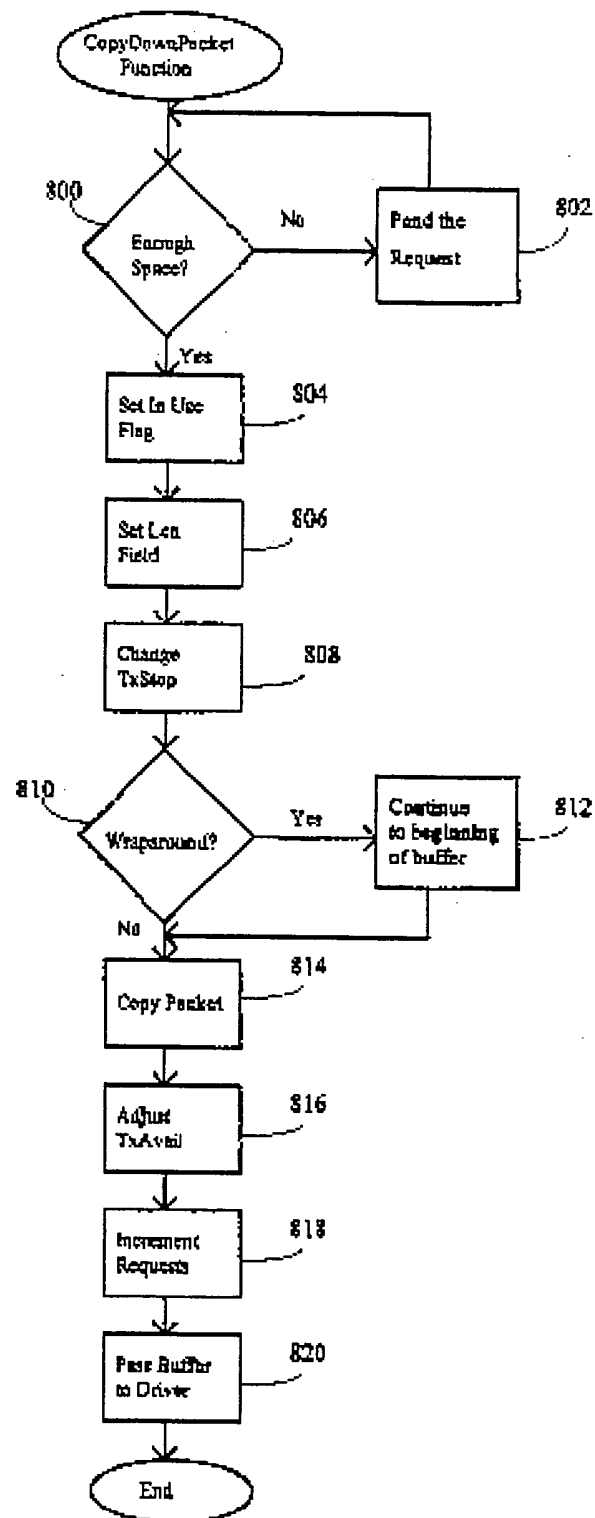


Figure 8

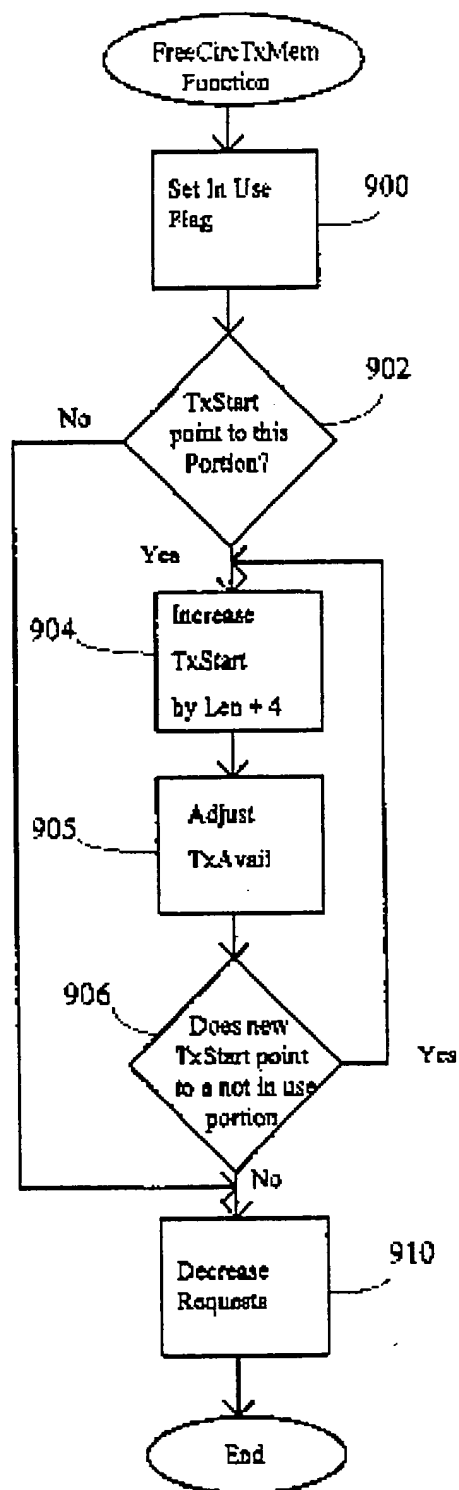


Figure 9

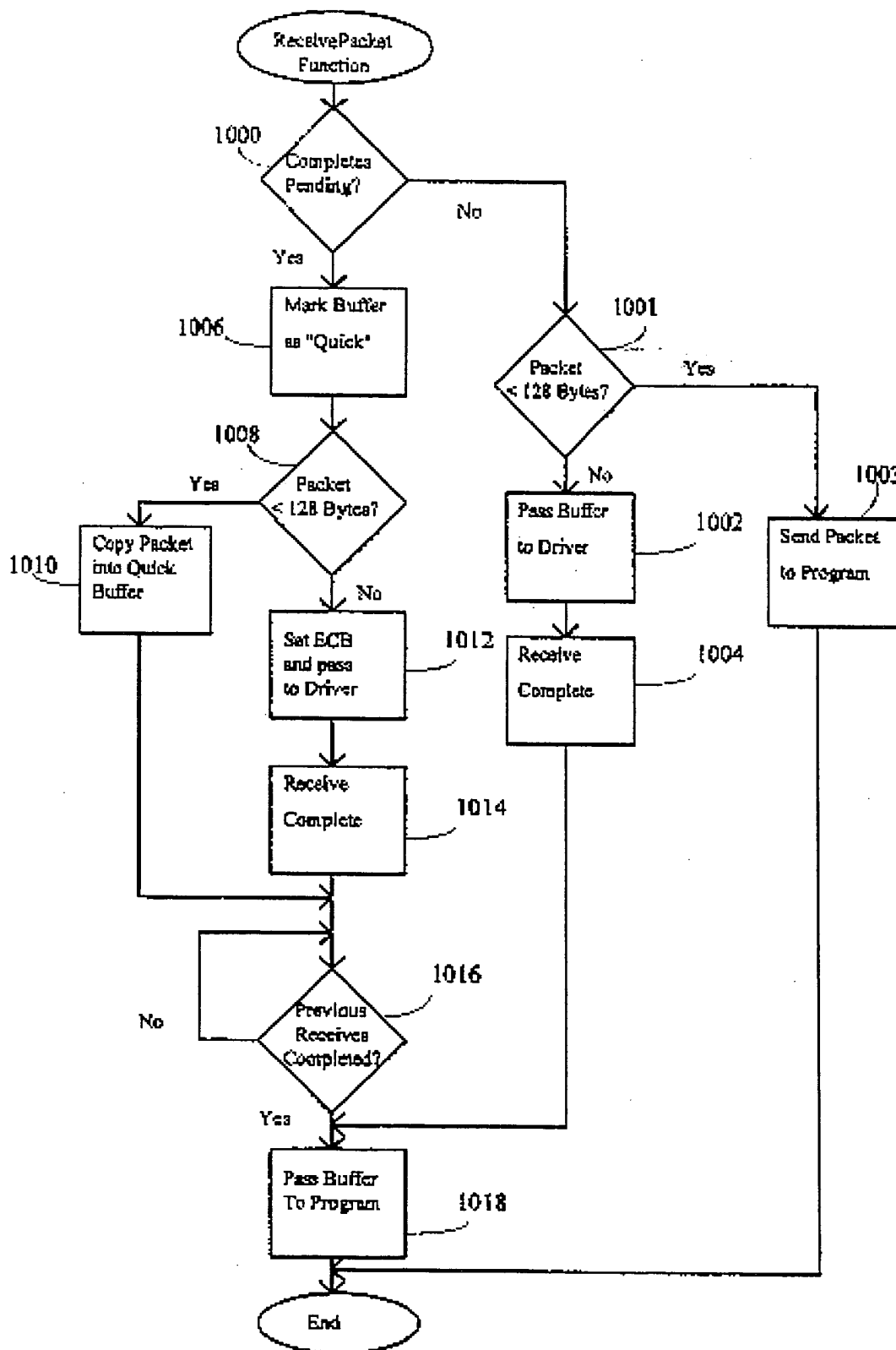


Figure 10

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) Publication number:

0 651 329 A3

(12)

EUROPEAN PATENT APPLICATION(21) Application number: **94116900.5**(51) Int. Cl.⁶: **G06F 9/46**(22) Date of filing: **26.10.94**(30) Priority: **28.10.93 US 144280**(43) Date of publication of application:
03.05.95 Bulletin 95/18(84) Designated Contracting States:
DE FR GB(86) Date of deferred publication of the search report:
30.08.95 Bulletin 95/35(71) Applicant: **MICROSOFT CORPORATION**
One Microsoft Way
Redmond,
Washington 98052-6399 (US)(72) Inventor: **Ogus, Aaron William**
9421 112th Avenue Northeast
Kirkland,
Washington 98033 (US)(74) Representative: **Patentanwälte Grünecker,**
Kinkeldey, Stockmair & Partner
Maximilianstrasse 58
D-80538 München (DE)(54) **Frame buffering of network packets.**

(57) An improved buffering mechanism for network packets is provided. In a preferred embodiment of the present invention, a circular buffer is used when sending and receiving network packets. The circular buffer is accessible by V86 mode network drivers and protected mode applications. Said circular buffer stores packets in a first-in-first-out manner, facilitates asynchronous communications across the network and services multiple clients.

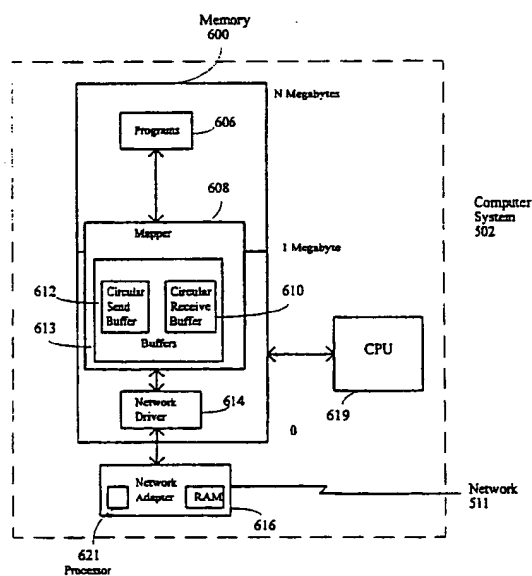


Figure 6



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 94 11 6900

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
A	IBM TECHNICAL DISCLOSURE BULLETIN., vol. 34, no. 10B, March 1992 NEW YORK US, pages 157-162, 'Multi-addressing Mode Microprocessor Memory Support for Personal Computers' * paragraph 2 - paragraph 3 * * paragraph 22 * ---	1	G06F9/46
A	PATENT ABSTRACTS OF JAPAN vol. 13 no. 210 (P-872) ,17 May 1989 & JP-A-01 026961 (ALPS ELECTRIC) * abstract * ---	1	
A	IBM TECHNICAL DISCLOSURE BULLETIN., vol. 34, no. 6, November 1991 NEW YORK US, pages 131-133, 'Direct Memory Access-Based Integrated etc...' * paragraph 1 * * paragraph 5 - paragraph 8; figures 1,2 * ---	1,2	
A	KNUTH 'The Art of Computer Programming, Volume 1: Fundamental Algorithms' 1988 , ADDISON-WESLEY , READING-US * page 236, line 13 - line 14 * * page 271, line 25 - line 26 * * page 437, line 1 - line 8 * ---	2	TECHNICAL FIELDS SEARCHED (Int.Cl.6) G06F
Y	IBM TECHNICAL DISCLOSURE BULLETIN, vol. 34, no. 11, April 1992 NEW YORK US, pages 340-350, 'Stimulating shared buffer communication etc...' * paragraph 9 - paragraph 16 * * paragraph 29 - line 12, paragraph 30 * --- -/--	15-18	
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 16 June 1995	Examiner Cohen, B
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ----- & : member of the same patent family, corresponding document	



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 94 11 6900

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Y	IBM TECHNICAL DISCLOSURE BULLETIN, vol. 26, no. 1, June 1983 NEW YORK US, pages 86-88, EVELAND ET AL 'Technique for storing variable length data etc...' * page 87, paragraph 2 - page 88, paragraph 2 *	15-18	
A	--- EP-A-0 198 559 (PHILIPS) * abstract * * page 7, line 10 - page 10, line 14; figure 2 *	15	
A	--- IBM TECHNICAL DISCLOSURE BULLETIN, vol. 28, no. 12, May 1986 NEW YORK US, pages 5314-5316, 'Software-Controlled Wrap-Around Buffer' * line 27 - line 49 * -----	15	
The present search report has been drawn up for all claims			TECHNICAL FIELDS SEARCHED (Int.Cl.6)
Place of search THE HAGUE		Date of completion of the search 16 June 1995	Examiner Cohen, B
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			



European Patent
Office

CLAIMS INCURRING FEES

The present European patent application comprised at the time of filing more than ten claims.

- ☐ All claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for all claims.
- ☐ Only part of the claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims and for those claims for which claims fees have been paid, namely claims:
- ☐ No claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims.

LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirement of unity of invention and relates to several inventions or groups of inventions.

namely:

See Sheet B.

- ☒ All further search fees have been paid within the fixed time limit. The present European search report has been drawn up for all claims.
- ☐ Only part of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the inventions in respect of which search fees have been paid, namely claims:
- ☐ None of the further search fees has been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the invention first mentioned in the claims, namely claims:



European Patent
Office

EP 94116900

LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirement of unity of invention and relates to several inventions or groups of inventions, namely:

Claims 1-14

Network computer with packet buffers in memory accessible by V86 mode.

Claims 15-18

Network computer with circular packet buffers.

THIS PAGE BLANK (USPTO)